Evolving Planning Behaviors in an Artificial Ecosystem

Baczewski, Joachim; Tani, Jun

Evolving Planning Behaviors in an Artificial Ecosystem

Joachim Baczewski
Department of Administration Engineering, Keio University
3-14-1 Hiyoshi, Kohoku-ku, Yokohama 223
Tel: 045-563-1141, Fax: 045-563-5979
jo@ae.keio.ac.jp

Jun Tani
Sony Computer Science Laboratory Inc.
Takanawa Muse Building
3-14-13 Higashi Gotanda, Shinagawa-ku, Tokyo 141
Tel: 03-5448-4380, Fax: 03-5448-4273
tani@csl.sony.co.jp

Abstract

A tradeoff between reactive and deliberative planning exists in artificial and natural ecosystems. We introduce an evolutionary approach to balance planning behavior for agents in competitive ecosystems. Agents are oriented towards their goal by forces which are designed in analogy to smell distribution in nature. Successful agents can have mutated offspring.

First, we explain our model and then we observe for different environments how the balance of deliberation and reactiveness evolves in our simulations. It can be seen that evolution adapts the planning behavior of agents successfully.

1 Introduction

In the competitive environment of natural and artificial real-time systems a tradeoff between the guarantee of execution of a plan and the cost minimization for the plan execution exists. We refer to this as a tradeoff between deliberativeness and reactiveness. This tradeoff is due to the *bounded rationality* [1] which characterizes living beings as well as intelligent agents in complex environments: optimizing its outcome with its limited abilities. We will focus on such problems in this paper by studying this tradeoff in an evolutionary system of multiple agents.

Evolution is a source for diversity and complexity, but also for stability in a dynamical system such as a system of multiple, autonomous agents which have to do planning in an ever changing dynamic environment in order to fulfill their goals. We assume such goal-driven agents to be selfishly motivated in the first place and to compete for resources. Executing a plan is connected with costs, e.g. energy consumed when moving a robot. Fulfilling a goal gives a payoff to an agent, e.g. reaching a loading-station for a robot. Even taking no actions is costly most of the time for the agents although on a much lower scale. This is obviously because we regard agents as dissipative systems [2] which consume energy in order to maintain their structure - as opposed to conservative systems. For a robot scenario this might be the energy consumed in order to do planning and to observe the environment.

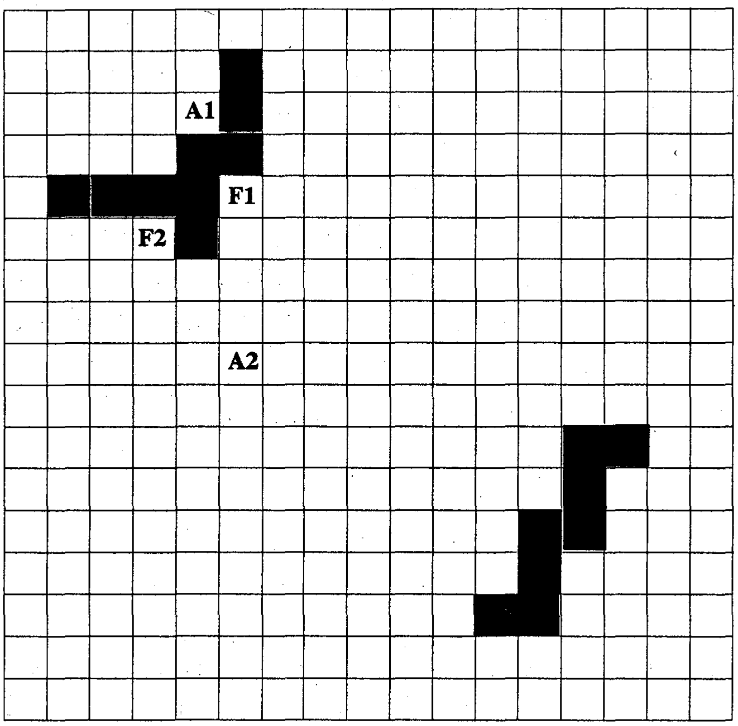

Figure 1: Tradeoff Scenario with Agents A1, A2 and Goals F1, F2, Obstacles.

Agents which reason deliberately can reduce arising costs by choosing a carefully selected low-cost plan. Such behavior represents a higher commitment to the environmental conditions. On the other hand, such committing agents might be confronted with a changed environmental situation when executing their plan after time-consuming deliberation in a dynamic environment [3]. This might result in failure of execution. Pure reactive agents will not spend much time on reasoning but will execute the fastest plan, which is normally not the cheapest, in order to fulfill their goal. Because reactive agents do not spend much or no time on planning they are likely to fulfill most of their goals as long as these goals are within their cost range.

Figure 1 clarifies the tradeoff which arises for agent A1. Assuming, because of his bounded rationality he has no knowledge about the existence of agent A2. But both agents can sense the goals F1 and F2. Fulfilling a goal is achieved by the agent which reaches it first and gives an immediate payoff the the agent. An ambivalent situation emerges because both agents will head for F1 which is closer then F2. The only chance for A1 to reach F1 before A2 would be to be perfectly reactive and to head for the goal disregarding the obstacle which is in between A1 and F1. This might be fatal if crossing the obstacle would consume all of A1's energy. On the other hand A2 will head for F1 as well and will fail to fulfill the goal if we assume A1 to be reactive. The term *obstacle* is used by us in a general sense and might as well refer to access speed to a resource or traffic on a network.

Examples for such scenarios from economy are buying and selling strategies on the stock market or marketing strategies of companies. Many such scenarios exist in biological ecosystems as well, e.g. hunting strategies of predators. Applications for evolutionary determined planning behavior of multiple agents might be found in computational market systems, where agents have to allocate resources or to retrieve information in distributed, open network configurations. Software agents in scheduling applications (e.g. power line or telecommunication line scheduling) have the same need for adaptive planning as multi robot systems.

In this paper we investigate how the balance of deliberativeness and reactiveness evolves in an artificial ecosystem. The unpredictability of a complex ecosystem increases the search space for a planning agent exponentially [4]. Thus, we assume planning to be much more time consuming for an agent then their actual acting. In general, while planning agents map their environment onto their planning model. The more sophisticated this model is, the lower will be the cost for the agent to execute the plan. The more sophisticated plan results in a longer time delay before acting. These time delays can reach critical values in a concurrent environment. In our study, the tradeoff between *time* and *goodness of plans* as a combinatorial optimization problem with normally exponential growing cost functions, is determined by a process of evolution.

Related work found in traditional AI includes the *Tileworld* [5] which was designed as a test-bed for agent planning behavior. In contrast to our work, an agent's planning characteristics in the Tileworld were neither autonomously determined nor emerging. This limits the usefulness for experiments in more complex dynamical environments. On the other hand, much research like *Artificial Ant* problems or artificial worlds like the *AL* world [6] can be found in the field of ALife. Such work focuses more on the process of evolution and development itself and might not be seen as a solution oriented approach for existing problems. Our approach tries to bridge the gap between those two worlds and the fields of AI/DAI and ALife by combining methods from both fields.

In the next section we will introduce how we model reactive and deliberative planning, what assumption we make about the ecosystem and what influence evolution has on the whole process. After that we will show the results for the evolution of planning behavior in our model for multiple agents. Finally we will conclude with discussions and give an outlook on related future research issues.

2 The Model

We chose a 2 dimensional discrete world as a simple search space. A number of agents is randomly placed in that world. Each agent has a limited sensing ability in all directions. There is no communication in this ecosystem between agents. Food and obstacles are placed in this world. It is each agent's goal to reach food. This is done by *mentally* moving within his sensing range. Mental as well as physical movement is possible in directions left/right/up/down. When food is found by this mental search process, the agent moves to the food, the food disappears and the agent gets an immediate payoff in form of energy increase. New food is created in fixed temporal steps which models natural growth processes.

2.1 Attraction - Repulsion

In order to model the attraction towards the food and the avoidance of obstacles in the world, we introduce two *forces* which are influencing each agent. These forces are determined in the following way in a discrete world where the agent can move in four directions.

Smell distribution in nature which is proportional to 1/radius in two dimensional space seems to be a suitable model for the attraction and repulsion modelling in an artificial ecosystem. We designed the attraction and repulsion force in analogy to the gradient of the *smell intensity*:

$$I(r) = c_1 I_0 / r \tag{1}$$

with $c_1 = constant$, I_0 =Intensity at $r = 0$.

Thus, the forces are designed according to:

$$gradI(r) = c_2 I_0 / r^2 \tag{2}$$

with $c_2 = constant$.

In Equation 3 to 6 the attracting force from a food-point to the agent is described. This is just a mental attracting force for the agent which focuses his interest on the food. r is the distance from the agent to the food. The closer the food is, the more it attracts the agent.

$$\vec{F}_{food} = (F_{food_r}, \Theta_{x,y}) \tag{3}$$

$$F_{food_r} = \begin{cases} F_0/r^2 & \text{if } r > 0 \\ 0 & \text{if } r = 0 \end{cases} \tag{4}$$

with

$$r = \sqrt{x^2 + y^2} \tag{5}$$

and

$$\Theta_{x,y} = \arctan(y/x) \tag{6}$$

In analogy, in Equation 7 to 10 the repulsion force is determined. This force can be seen as a real physical force for the agent considering that, e.g., the closer a robot moves to a wall the stronger is the repulsion, the more difficult it gets for him to navigate.

$$\vec{F}_{obst} = (F_{obst_r}, \phi_{x,y}) \tag{7}$$

$$F_{obst_r} = \begin{cases} O_0/r^2 & \text{if } r > 0 \\ 0 & \text{if } r = 0 \end{cases} \tag{8}$$

with

$$r = \sqrt{x^2 + y^2} \tag{9}$$

and

$$\phi_{x,y} = \arctan(y/x) \tag{10}$$

F_0 and O_0 are food and obstacle constants. We defined the forces to be 0 for the singular point $r = 0$ because food, on the one hand, is consumed when the agent reaches the food location and the attraction stops. On the other hand, when an agent *climbed* an obstacle there will be no force on him from that obstacle.

The total forces on an agent are listed in Equation 11. The factor k is a biasing factor which makes the agent more deliberative for small k by emphasizing the repulsion forces and more reactive for high k by emphasizing the food attraction. The factor k is determined evolutionarily by mutation when creating offspring.

$$\vec{F}_{agent} = \sum_j \vec{F}_{obst_j} - k \sum_i \vec{F}_{food_i} \tag{11}$$

Because we are using a discrete space model it is inconvenient to deal with force vectors directly. Thus, we integrate the forces into a *potential field.* On this potential field the agent's decisions are based. The agent sees himself located in a potential field landscape and moves mentally to the lower potential. The bias factor k determines the steepness of the attracting food gradient.

2.2 Search

First an agent searches, starting from his location, with his mental focus-point within his sensing range. The search is determined according to the potential of the adjacent fields of the position of the mental search process. The search process moves always to lower potentials. After finding food the agent moves to the food position. The time required for planning steps is longer then the moving time. If two agents have the same food as goal, the faster agent gets the food and the other agent stops the planning process. Every time step the agent's energy decreases gradually. Moving over an obstacle results in a penalty.

The initial fitness of an agent increases by

- consuming food

and decreases

- by moving over an obstacle (penalty)
- as a linear function of time,
- when creating offspring.

As in our previous work [7] we apply a backtracking mechanism based on that introduced by Ishida [8]. An agent gets to know when he reached food. Thus, it can be decided easily that on a location without food but with lower potential than the potential of the adjacent fields a local minimum was reached. In that case the agent increases the potential of that point step by step until he is able to leave this location. Loops in agent plans are cut.

In order to create offspring an agent has to reach a certain energy level. Offspring inherits its parent's characteristics which undergo mutations.

Our mutation operators mutate

- the bias factor k
- the size of the mental potential used in local minima avoidance
- the mutation rate itself

in small steps which ascertains a smooth drift to more optimal parameters but which allows in a open environment the emergence and development of a variety of planning strategies. Mutation of the mutation rate allows the system to find its optimal parameters autonomously [9]. The offspring is situated randomly in the environment in order to avoid clustering. Existing elite in the environment is preserved because the parent still exists after reproducing. If an agent's fitness is lower than a certain minimum value the agent dies and disappears.

3 Experiments

We were interested in the development of the factor k in an ecosystem of multiple competitive agents. In order to specify the influence of the environmental difficulty, we observed the average k in the system in relation to the maximum k which is given as a value for which nearly perfect reactivity is achieved. Various simulations showed us that the ecosystem is very sensitive to parameter-settings. For certain parameter-settings the agents die out after a while. For the simulations shown here we chose a parameter range where after a number of generations in this competitive world a detailed balance between agent and ecosystem emerged.

Figure 2 shows a simple obstacle constellation. A number of agent and a number of food was located in this scenario. In Figure 3 the development of k/k_{max} is shown for this scenario.

Figure 4 shows a quite difficult obstacle constellation. A number of agents and a number of food was located in this scenario, too. In Figure 5 the development of k/k_{max} is shown for this scenario.

Agents plan as feedback to the environmental conditions as can be seen from our results. Figure 3 shows a tendency towards favoring more reactive agents because the force field in scenario 1 allows agents to be reactive to a certain extend without getting a penalty unless food is located behind an obstacle.

Figure 5 shows a sudden drift towards deliberative planning. If agent A in Figure 4 would be a reactive agent he would cross the obstacle in order to get to the food. Because this scenario has some difficult obstacle constellations reactive agents frequently cross over obstacles and use much of their energy until they get extinct. At that point deliberative agents take over and settle down around $k/k_{max} = 0.2$.

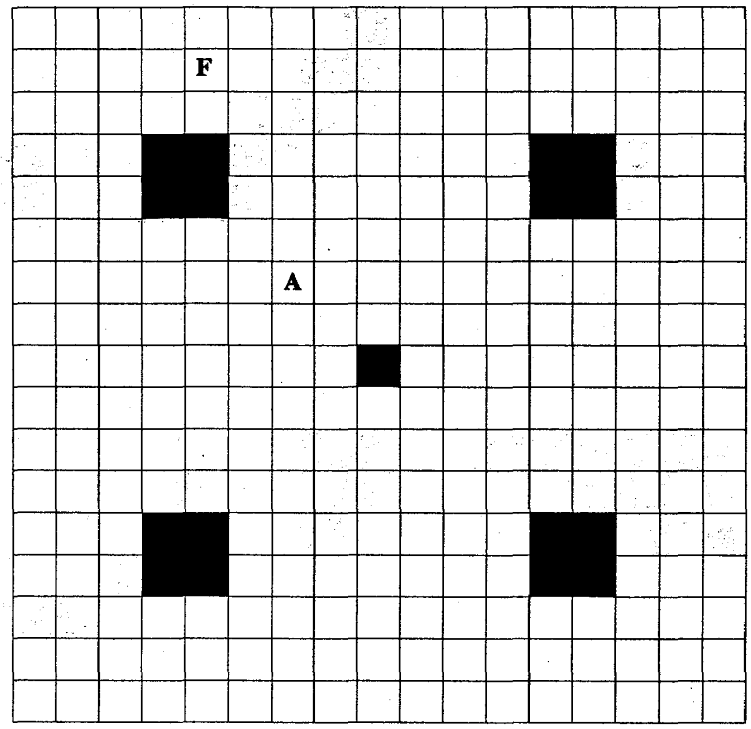

Figure 2: Scenario 1 (with Agent A and Food F as example).

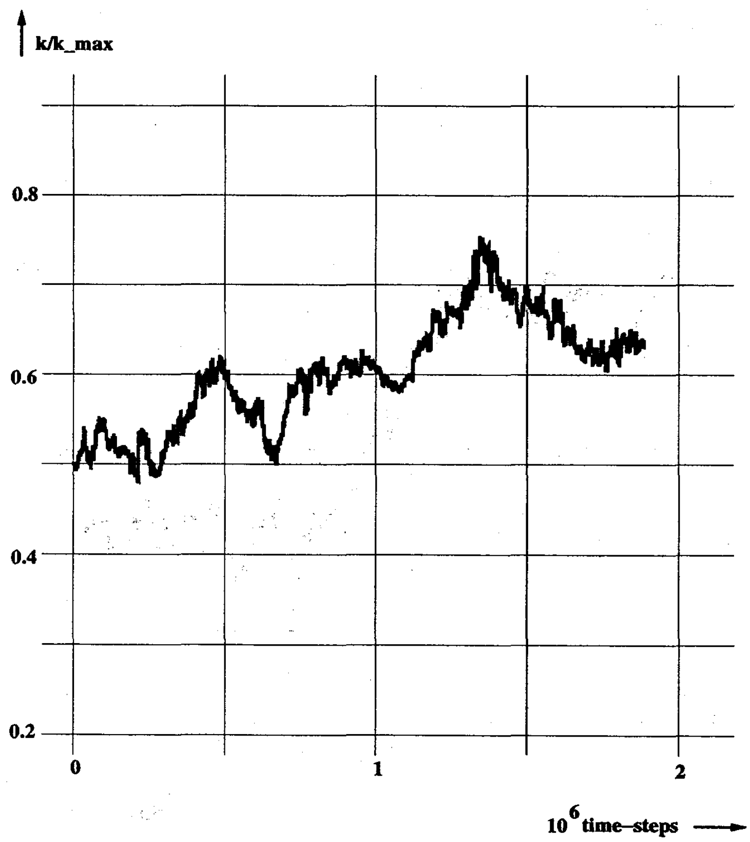

Figure 3: Development of average k/k_{max} for Scenario 1.

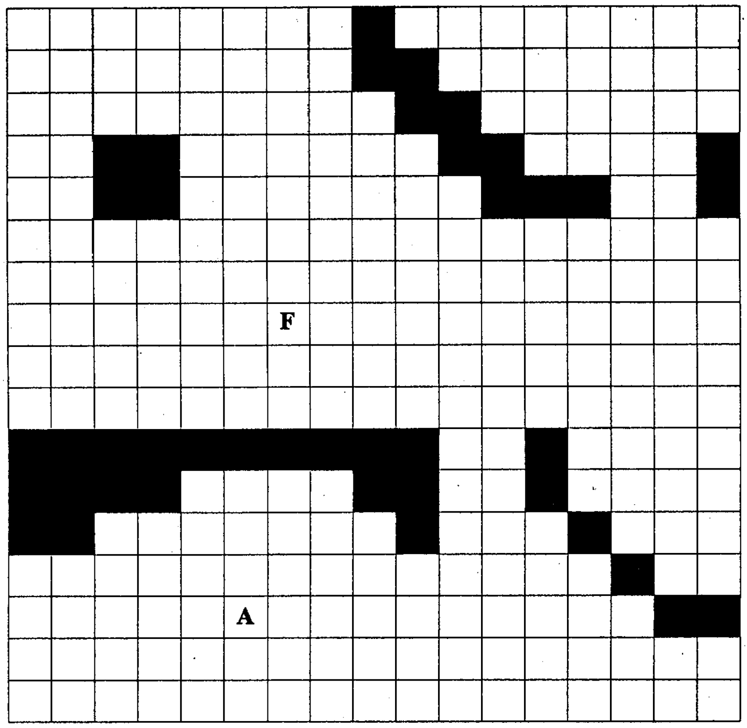

Figure 4: Scenario 2 (with Agent A and Food F as example).

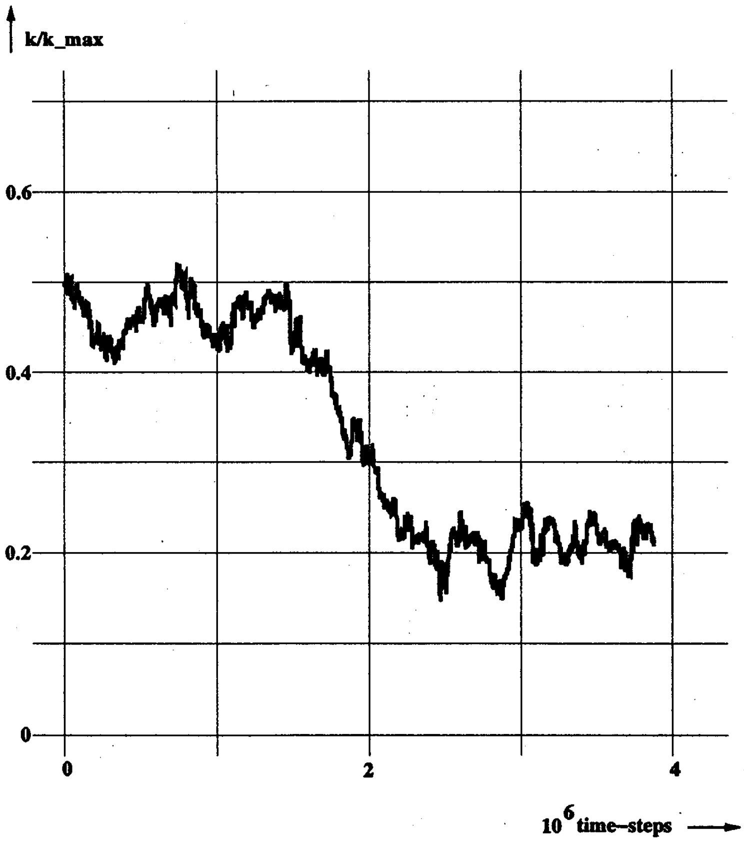

Figure 5: Development of average k/k_{max} for Scenario 2.

4 Discussion and Conclusion

In competitive ecosystems, planning needs to be balanced between deliberative and reactive planning. Maladapted goal-oriented agents in multi-agent systems will not perform well. Thus, there is the need for adaptive planning behavior. We showed that with a smell distribution-like modeled potential field method and mutation operators evolution can find a balance for various environmental difficulties. Adaption is the key to the solution for finding optimal planning behavior of agents.

For future work we are interested in specifying the emerging planning diversity more clearly. We further like to observe how the stability of emerging interactions depends on the mutation rate. A changed mutation rate can increase the learning of the right balance between deliberativeness and reactiveness but can in some cases also bring unstable behavior which can lead to a ecosystem breakdown.

References

[1] Stuart Russell. Rationality and Intelligence. In *Proceedings of IJCAI 95*, pp. 950-957, 1995.

[2] Grégoire Nicolis and Ilya Prigogine. Exploring Complexity - An Introduction, Freeman, 1989.

[3] David N. Kinny and Michael P. Georgeff. Commitment and Effectiveness of Situated Agents. In *Proceedings of IJCAI 91*, Volume 1, pp. 82-88, 1991.

[4] Philip E. Agre and David Chapman. Pengi: An Implementation of a Theory of Activity. In *Proceedings of AAAI 87*, Volume 1, pp. 268-272, 1987.

[5] Martha E. Pollack and Marc Ringuette. Introducing the Tileworld: Experimentally Evaluating Agent Architectures. In *Proceedings of AAAI 90*, Volume 1, pp. 183-189, 1990.

[6] David Ackley and Michael Littman. Interactions between Learning and Evolution. In *Artificial Life 2*, pp. 487-509, 1992.

[7] Joachim Baczewski and Jun Tani. Evolutionary Agent Planning. In *Proceedings of ICEC 96*, pp. 744-747, 1996.

[8] Toru Ishida. Moving Target Search with Intelligence. In *Proceedings of AAAI 92*, pp. 525-532, 1992.

[9] Kunihiko Kaneko and Takashi Ikegami. Homeochaos: dynamic stability of a symbiotic network with population dynamics and evolving mutation rates. In *Physica D 56*, pp. 406-429, 1992.